Recent Developments in Population Projection Methodology

Article *in* Population Space and Place · September 2005

DOI: 10.1002/psp.389 · Source: OAI

2 authors:

Tom Wilson

Philip Howell Rees

University of Leeds

POPULATION, SPACE AND PLACE
Popul. Space Place **11**, 337–360 (2005)
Published online in Wiley InterScience (www.interscience.wiley.com). DOI: 10.1002/psp.389

Recent Developments in Population Projection Methodology: A Review

Tom Wilson[1]* and Phil Rees[2]

[1] *Queensland Centre for Population Research, School of Geography, Planning and Architecture, University of Queensland, Brisbane, Qld 4072, Australia*
[2] *School of Geography, University of Leeds, Leeds LS2 9JT, UK*

ABSTRACT

In this paper we survey five streams of research that have made important contributions to population projection methodology over the last decade. These are: (i) the evaluation of population forecasts; (ii) probabilistic methods; (iii) experiments in the projection of migration; (iv) projecting dimensions additional to age, sex and region; and (v) the use of scenarios for 'what if?' analyses and understanding population dynamics. Key developments in these areas are discussed, and a number of opportunities for further research are identified. Copyright © 2005 John Wiley & Sons, Ltd.

Received 17 March 2005; revised 20 June 2005; accepted 20 June 2005

Keywords: population projections; review; methodology; probabilistic projections; migration; scenarios

INTRODUCTION

A key competence of successful countries is the ability to plan flexibly for the short and long term. Planning must be based on good information about the present situation, on how important variables will change in the future and on how much faith to place in future predictions. One significant variable that enters national and local planning is population. Much effort is expended by researchers and governments in projecting the future of the population under a given set of assumptions. Often alternative projections are made, and from these a forecast is chosen upon which to base planning.

To combat uncertainty, projections are frequently refreshed (often every year for national populations, less frequently for subnational projections) using the most recent information about demographic trends. To provide guidance about uncertainty, statistical offices develop variant projections based on alternative scenarios about the projection 'drivers'. More recently researchers have attempted to be more precise about uncertainty and have made estimates of the probability of outcomes. Population projection, despite its uncertainties, is important for our futures in pension planning by the state, in preparing health provision, in delivering future retail services and in improving the education of the population. This is why research into population projection methods and assumptions can be found in such a wide variety of social science disciplines. Research teams on larger projects are often drawn from a wide range of disciplines.

Over the years, through interaction between the suppliers and users of demographic data, the scope and accuracy of information input to projections has improved. The greatest uncertainties concern the migration component of projections; at the national scale, international migration estimates are subject to error, while for regions, internal migration data series suffer from gaps in time coverage or discontinuities in geographical coverage. One technical development has made possible new and better analyses: the increase in computer speeds, processor memory and hard drive capacity.

*Correspondence to: T. Wilson, School for Social and Policy Research, Charles Darwin University, Darwin, Northern Territory 0909, Australia. E-mail: tom.wilson@cdu.edu.au

It is against this background that we carry out our review. All reviews of literature are necessarily selective. Here we mostly limit ourselves to work published in English subsequent to the comprehensive review by Willekens (1990). We assume the reader is familiar with the main cohort-component method of population projection and its various implementations prior to 1990. In addition, rather than attempting the wide coverage of Willekens, we focus on five themes that have been important in the advancing population projection methods, namely:

(i) *evaluation of population forecasts* ('what can we learn from projection errors?')
(ii) *probabilistic methods* ('how can we define the range of uncertainty?')
(iii) *migration projections* ('how can we project population movement?')
(iv) *projecting other dimensions* additional to the demographic standards of age, sex and region ('how can projection models be employed to address new issues?')
(v) *scenarios* ('what alternative futures should we explore?').

We will say relatively little about fertility and mortality projection, developments in mathematical demography and small area projections, which could be the subject of separate reviews. Before we embark on our five-fold review we provide a brief historical survey to place recent work in context.

A BRIEF HISTORY OF POPULATION PROJECTION METHODS

Over the twentieth century, population projections changed from being seen as the fitting of deterministic mathematical functions to the total population, to one of articulating the multiple processes of demographic change effected through appropriate population groupings, which experience the processes at different intensities (de Gans, 1999). The cohort-component model of Whelpton (1936) had been formalised by mid-century in matrix form (Leslie, 1945, 1948) and has been used ever since in country projections by national (e.g. GAD, 2004; Shaw, 2004) and international statistical organisations (e.g. United Nations, 2004).

The period between 1966 and 1986 saw the expansion of the cohort-component model to deal with multiple interacting populations: Rogers (1966) introduced the multiregional model, linked it to the life table based cohort-component model (Rogers, 1975), developed a parameterised version (Rogers, 1986), and with Willekens organised its application to many countries (Rogers and Willekens, 1986). Parallel work on the multistate model was carried out by Schoen (1988). At the start of the next decade van Imhoff (1990) showed how the multistate model could be defined assuming exponential rather than linear changes in demographic intensities by age.

The detailed review by Willekens (1990) outlined the development of demographic projections over the 1970s and 1980s. It is useful to use his framework to provide a context for more recent research. Willekens outlined a conceptual framework in which projections are organised in eight steps (Cruijsen and Keilman, 1984: cited in Willekens, 1990): (a) system identification ('what populations, groups, time intervals are we interested in?'); (b) system description ('what's been happening in the recent past?'); (c) model design ('what kind of projection model is suitable and feasible?'); (d) hypothesis formulation ('how do we think fertility, mortality and migration are changing and under what assumptions?'); (e) sensitivity analysis ('which assumptions really make a difference to outcomes?'); (f) implementation ('what software do we have for running our projections?'); and (g) monitoring ('how did we do in our projections now that the future is the past?'). Our discussion of evaluation falls into step (g); our discussion of probabilistic methods can be placed in step (e); migration projections are part of step (d); additional dimensions used in projection relate to steps (a), (b) and (c); scenarios are one form of step (d).

Willekens distinguished models using extrapolation of the projection drivers from models attempting explanation of the projection drivers. But he argued strongly that when 'the weather is changing', extrapolative models perform poorly and explanatory models will do better. Most current population projections belong to the former class and generally perform better than explanatory models (Wilson *et al.*, 2002; van der Gaag *et al.*, 2003a, b) when comparative tests are performed. However, those tests have often been conducted in a 'stable weather period' (the low fertility, low and improving mortality and high but fluctuating immigration regime of the last decade in Europe). This tension between description and explanation in population projections is likely to continue.

THE EVALUATION OF POPULATION FORECASTS

What does Evaluation Involve?

The literature on population forecast evaluation appears to suggest two interpretations of 'evaluation': a narrower definition which refers only to forecast accuracy, and a broader one which encompasses 'provision of necessary detail, face validity, plausibility, costs of production, timeliness, ease of application and explanation, usefulness as an analytical tool, political acceptability, and forecast accuracy' (Smith *et al*., 2001: 279). This broader definition therefore includes not just *ex post* evaluation of forecast figures, assumptions and model suitability, but also addresses questions asked by forecasters such as 'Which model should I use?' and 'How should I go about formulating the forecast assumptions?'. In contrast to population estimation (e.g. Simpson, 1998) there is little guidance for practitioners on model selection or assumption setting, and what does exist is specialised and scattered in the literature. Here we restrict our focus to the narrower definition of forecast evaluation and consider a number of recent studies of forecast accuracy (see Chapter 13 of Smith *et al*., 2001, on measures of forecast accuracy). The term '*forecast* accuracy' rather than 'projection accuracy' is deliberately used here. Population projections are statements about the future and may be likely, plausible or impossible; population forecasts are statements about the most likely demographic future in the view of the analyst. Projections are always correct (unless there are computational errors) and so an evaluation of their accuracy is nonsensical; forecasts are nearly always wrong to a degree.

Errors in National and Global Forecasts

Keilman (1997) reviewed the literature on national forecast errors for selected industrialised countries and provided detailed evaluations for Norway and the Netherlands. This study extended earlier work by examining dependency ratio forecast errors, and also whether errors in births and deaths are correlated (something which would be useful to know in probabilistic forecasting). The literature review reveals a fairly common pattern of modest forecast errors for most adult age groups, but large errors at both ends of the age spectrum, with over-projections amongst children and under-projections at the oldest ages – due to over-optimistic fertility and pessimistic mortality assumptions. In Keilman (1999, 2001) the spotlight shifted to the United Nations' forecasts for the world, global regions and large countries. Amongst the findings were: total population forecasts for the world are highly respectable; improvements in accuracy can be seen over time; considerable differences are evident in forecast accuracy from region to region; and age-specific errors are generally greatest at the youngest and oldest ages.

Errors in Subnational Forecasts

In their evaluations of US population forecasts, Smith and Shahidullah (1995) and Smith and Tayman (2003) extended the analysis of forecast errors by age group to subnational areas. The addition of internal migration into the equation resulted in age patterns of error with some differences from those at the national level. At the state level the largest absolute percentage errors were found at the youngest ages and in the peak migration ages of the 20s and 30s. County-level forecast errors by age were calculated for Florida. These exhibited less of an identifiable pattern of error across age groups, with different error patterns in forecasts with different jump-off years.

An evaluation of subnational forecasts produced by consultants for the Directorate General for Regional Policy was carried out for the European Commission by Rees *et al.* (2001). The evaluation reviewed methods used in the four forecasting rounds completed by 1998, and recommendations for further improvements were made. The evaluation reviewed errors in populations and components by comparing forecast populations with population estimates generated subsequent to the forecasts. This evaluation showed that errors were reduced through successive rounds, mainly because of the addition of better migration variables and inputs. Projections based in 1980 and 1985 failed to include international migration at all. The 1990 and 1995-based forecasts underestimated the flow of international migration into the European Union. The childhood ages were less well projected and the retirement ages systematically under-predicted. Examination of the sequence of regional forecast populations for the UK suggested important shifts between the 1980 and 1985 rounds, which forecast rapid decline for larger cities, and those for 1990 and 1995, which forecast slower decreases, stability or even some population increase.

The Errors of Hypothetical Forecasts

Rather than assessing actual past forecasts, interesting results have been obtained by carrying out hypothetical 'forecasts' from some historical jump-off date and comparing with actual data (an approach which has parallels with Alho's volatility-based method mentioned in the next section). We give two examples here. Tayman and Swanson (1996) asked: 'how much added value is provided by a cohort-component or land-use based forecast compared with a naïve one of no change?' This approach thus complements the complex versus simple models debate in *Mathematical Population Studies* 5.3 (1995). These authors calculated the 'proportionate reduction in error', the amount by which forecast error is reduced by using the more complex method rather than the naïve one. In many US county and census tract forecasts the naïve method often performed better!

The second study was carried out as part of the National Research Council's (2000: Appendix E) research on forecasting the world's population. Three methods for forecasting international migration for countries of the world were implemented and compared with UN estimates for the selected period. The first method assumed no migration; in the second, net migration was held constant from some base period; and the third method consisted of constant net immigration for industrialised countries, with this migration distributed as 'net emigrants' to developing countries in proportion to population. It was found that for forecast horizons of longer than ten years, the mean absolute error was greater for the constant migration forecasts than those which excluded migration. This is an average over all countries, however. For high net immigration countries such as Australia, Canada and the US, the 'no migration' approach proved inferior. For the third approach the mean absolute error in net migration was fairly similar to the 'no migration' simulations.

The Nature of Forecast Errors

The statistical properties of forecast error measures have attracted attention. Tayman and Swanson (1999) highlighted the tendency of the mean absolute percentage error (MAPE) distribution to be skewed due to outliers, thereby exaggerating average error. They examined symmetrical MAPE and several types of M-estimators as alternatives to MAPE, suggesting that two of these M-estimators would be useful additions to the toolkit of forecast accuracy measures. However, whilst giving less biased indications of forecast error, these M-estimators are difficult to interpret. In a subsequent paper, therefore, Tayman *et al.* (1999) went on to examine ways in which the MAPE distribution might be transformed to reduce its skewness. They proposed a nonlinear transformation of MAPE which was then re-expressed in the original scale of MAPE for ease of interpretation. We note that there is, of course, a large body of statistical literature on measures for comparing variables available for the evaluation of population forecasts.

PROBABILISTIC METHODS

The past decade has seen demographers developing methods to deal with error explicitly – this is undoubtedly the most important development in the period.

Uncertainty in Population Forecasting

Although the first papers on probabilistic population forecasts appeared in the 1980s, recent years have seen many methodological advances and an increase in the number of applications to countries and regions. The rationale behind probabilistic methods is that population forecasting is unavoidably uncertain. Whilst forecasts of total populations over short time periods have been reasonably reliable, forecasts for certain age groups – especially the oldest and youngest – have often been wide of the mark, as have forecasts of fertility, mortality and migration. Such uncertainty stems from many sources, including a limited understanding of the forces behind fertility, mortality and migration, the inherent tendency of observed demographic values to vary randomly on either side of a trend, and errors in measuring past demographic trends and the jump-off population. A detailed consideration of these and other issues surrounding uncertainty in population forecasting is provided by de Beer (2000).

But while experience has demonstrated the difficulty in generating precise population forecasts, methods have been devised for the next best thing: to estimate the extent to which forecasts might turn out to be wrong. Rather than providing a single forecast population number for each year in the future as is the case with conventional deterministic projections, probabilistic forecasts give a range of numbers and the estimated probability that the population will be within that range. The basic idea is as follows. Thousands of different trajectories of future age-specific rates are prepared. These are created by randomly selecting key trajectory parameters within specified error distributions. A cohort component model is then run thousands of times with a different set of varying age-specific rates input to each run. The result is a large set of outcome variables (population, births, deaths, migrations) for each future year. These are sorted into size order from smallest to largest and cumulative probability distributions of outcomes computed. For example, Lutz *et al.* (2001) have calculated that there is an 80% probability that the world's population will lie between 7.4 and 10.4 billion by 2050, with a median figure of 8.8 billion.

The Problems with High, Medium and Low Variants

Traditionally, the uncertainty of population forecasting has been handled by the production of 'uncertainty variants', typically labelled high, medium and low. This approach has been criticised in recent probabilistic work. Most users will understandably take the medium projection as the most likely population future, particularly as the meaning of the 'high' and 'low' variants is rather ambiguous. Just how likely is it that population will follow the high or low paths? Are they highly unlikely, or quite probable alternative population futures? (Lutz and Scherbov, 1998). There are a number of other problems with this deterministic variants approach. As Lee (1999) pointed out, uncertainty variants are probabilistically inconsistent. The high–low range for one demographic variable (such as total population size) does not cover the same probability as another variable (such as the elderly dependency ratio). Furthermore, the future trajectories of fertility, mortality and migration are normally set to linear or smooth curvilinear paths. This rules out random fluctuation or cyclical paths. For details of the conceptual problems of uncertainty variants, see, for example, de Beer (2000), Keilman *et al.* (2001) and Lee (1999).

Approaches to Quantifying Uncertainty

A variety of methods has been proposed for generating predictive intervals for population forecasts, and a number of authors have distinguished three broad approaches: (a) model-based estimates; (b) expert judgment; and (c) *ex post* methods – i.e. the use of past forecast errors (National Research Council, 2000; Alders and de Beer, 2004). These are considered in turn.

The first approach involves fitting autoregressive integrated moving average (ARIMA) models to fertility, mortality and migration trends. A key strength of ARIMA modelling is that forecasters can draw upon a large body of existing statistical theory and software. Some forecasters fit these models directly to summary demographic drivers such as the total fertility rate (e.g. Lee, 1993), whilst others have calibrated them to the parameters of a mathematical curve of age-specific rates (such as Keilman and Pham, 2000). For mortality forecasting the Lee-Carter method has proved popular (Lee and Carter, 1992). For examples of applications, see Booth *et al.* (2002), Lee (2000), Renshaw and Haberman (2003), Tuljapurkar *et al.* (2000), Wilmoth (1996) and Wolf (2004).

Expert judgment approaches are most commonly associated with the work of Lutz and colleagues. In their earlier probabilistic work (such as Lutz *et al.*, 1996) experts were asked to provide a 90% predictive interval (the 5% and 95% confidence levels) for fertility, mortality and migration for a certain period in the forecast horizon. Linear paths for each of these three demographic indicators were drawn at random from the distributions. The linearity of these paths was criticised by Lee (1999) for failing to allow cyclical and annual random variations, and in later work Lutz and colleagues have incorporated this randomness (e.g. Lutz *et al*., 2001a). Some researchers have also questioned the use of expert judgment in setting the widths of the predictive intervals (Lee, 1999). However, this approach permits the input of demographic knowledge into the forecasting process, whereas time series models use only past data to forecast the future (Lutz *et al*., 1996). The approach used by Lutz and colleagues has evolved into what they term expert argument-based probabilistic forecasting, details of which are given in their book *The End of World Population Growth in the 21st Century* (Lutz et al., 2004).

A rather different approach to preparing predictive intervals is to assume that the distribution of forecast errors calculated for past forecasts provides a valid estimate of forecast errors in the future. This is termed the *ex post* method. An early contribution is due to Stoto (1983) who produced predictive intervals for official US population forecasts. Recently the National Research Council (2000) published probabilistic population forecasts for world regions and selected countries from a model based on this approach. Alho and Spencer (1997) extended this idea using what they term volatility-based methods. In this method, naïve forecasts of demographic variables (where the values are held constant or assumed to change constantly over time) are produced from some jump-off year in the past and the forecast errors are calculated. The advantage of this approach is that it can be applied in situations when there are too few past sets of forecasts to give reliable distributions of errors.

In fact, the three approaches of constructing predictive intervals listed above are not mutually exclusive, and the most recent probabilistic forecasts tend to use a combination of at least two of these approaches. For example, Keilman *et al.* (2001) constrained their life expectancy at birth forecasts to the official Statistics Norway medium assumptions, and de Beer and Alders (1999) estimated the predictive interval for the total fertility rate (TFR) which is informed by both a time series model and a parity-specific cohort analysis of fertility. One could argue that all probabilistic approaches incorporate (expert) judgment in one form or another, explicitly or implicitly (Lutz *et al*., 2001a). The length of the time period used to fit an ARIMA model is an important decision that influences the width and level of the predictive intervals. For example, for fertility, does one fit the model to the last 60 years of data which includes the baby boom, or just the more recent 30-year period of below-replacement fertility?

Applications

To date, the countries for which probabilistic forecasts have been produced include Australia (Wilson and Bell, 2004a), Austria (Lutz and Scherbov, 1998), Finland (Alho, 2002a), India (Lutz and Scherbov, 2004b), Lithuania (Alho, 2002b), the Netherlands (de Beer and Alders, 1999), New Zealand (Wilson, 2004), Norway (Keilman *et al*., 2001), Singapore (Lutz and Scherbov, 2004a), South Africa (Lutz *et al*., 2001b), the US (Lee and Tuljapurkar, 1994) and the UK (Coleman and Scherbov, 2005). Lutz and his coworkers have also produced probabilistic forecasts for the European Union (Lutz and Scherbov, 1999) and world regions (Lutz *et al*., 2001a; Lutz *et al*., 2003a, b; Sanderson *et al*., 2004), as well as the world as a whole (Lutz *et al*., 1997). Probabilistic global population forecasts have also been prepared by Alho (1997). In addition, the National Research Council Panel on Population Projections (2000) has published probabilistic population forecasts for world regions and selected countries. The panel used a statistical model based on past UN population forecast errors.

One other example, arguably the most ambitious probabilistic forecasting exercise to date, is the Uncertain Population of Europe (UPE) project, the aim of which is to develop probabilistic population forecasts for 18 countries of the European Economic Area (the EU15 prior to May 2004 together with Iceland, Norway and Switzerland) (Statistics Finland, 2004; Alho and Nikander, 2004; Keilman and Pham, 2004). A combination of methods was used to prepare predictive intervals for fertility, mortality and migration. Heering and Keilman (2004) report experiments in which a few international experts were asked to choose 10% and 90% confidence levels for key population drivers (life expectancy at birth, total fertility rate and net international migration). The experts were asked to choose point forecast values and intervals choosing the naïve model, the empirical model, the ARIMA model, or to suggest alternatives. The naïve intervals were based on a constant projection of the driver variable, the empirical intervals were based on an analysis of historical errors, and the ARIMA intervals were generated from use of that model with suitable time series.

Extending Probabilistic Methods

In recent years a few authors have begun to consider how probabilistic methods might be extended to dimensions other than age and sex, including households (Leiwen and O'Neill, 2004), state finances (Lee and Tuljapurkar, 1998), greenhouse gas emissions (O'Neill, 2004) and subnational areas. Here we focus in particular on subnational forecasts. The addition of internal migration to population forecasting uncertainty presents additional challenges, since time series of past migration trends are frequently shorter than at the national level, less detailed, and sometimes of dubious quality. Internal migration tends to fluctuate over time in response to cyclical economic and housing factors, rendering it difficult to forecast. This greater uncertainty of the demographic future at subnational scale is why a probabilistic approach is even more important than at the national level.

Possibly the first contribution to subnational probabilistic forecasting was that of Rees and Turton (1998). They applied a multiregional method and judgment-based predictive intervals to present probabilistic forecasts for 71 regions of 12 European Union countries. Wilson and Bell (2004b) produced probabilistic population forecasts for Queensland and the rest of Australia in which predictive intervals were specified for internal migration gross migraproduction rates. Recently, Gullickson (2001) has emphasised the benefits of marrying the knowledge gained from national-level probabilistic forecasting with multiregional demography, and suggested that a log-linear modelling framework would be a useful way of forecasting the internal migration matrix.

Other subnational population forecasts have employed single regional methods. Gullickson and Moen (2001) presented a probabilistic forecast for two regions of Minnesota in order to predict the number of trauma patients for a large hospital in Minneapolis. These authors used the Lee (1993) method for forecasting fertility and the Lee-Carter method for mortality forecasts (Lee and Carter, 1992). Due to data limitations they used a simple procedure for estimating predictive intervals for crude net migration rates together with a deterministic age pattern. More recently, Miller (2002) and Lee *et al.* (2003) prepared probabilistic population and fiscal forecasts for California. Internal migration was handled by specifying predictive intervals for net internal migration, whilst for international migration the US Census Bureau medium variant net international migration at the national level was taken as given, and a predictive distribution of the percentage in California was prepared.

Whilst the cohort component model has been dominant in probabilistic population forecasting, some researchers have been interested only in forecasting total populations and have used simpler methods. Smith and Tayman (2004) have experimented with the use of ARIMA models to forecast the population totals of the US states directly. Using a variety of forecast periods and types of ARIMA model, the authors conclude:

> 'we are not convinced that – at this time – confidence intervals based on time series models can provide a realistic indication of the uncertainty associated with state population forecasts.' (p. 20)

Further research is needed to determine whether this approach can be modified into a reliable method.

An alternative approach to creating subnational predictive intervals is to generate a point forecast and apply predictive intervals using an *ex post* method. Smith and Sincich (1988) conducted forecasting experiments by applying a series of simple models to project the total population of US states over different time periods. They discovered that sufficient temporal stability existed in population forecast errors for them to be used as a reliable guide to likely future errors. For small area population forecasts, however, a sufficient number of past forecasts is unlikely to be available. In addition, small areas tend to be subject to regular boundary changes which would invalidate the method. The solution to these difficulties proposed by Tayman *et al.* (1998) is to obtain a large sample of forecast errors from a limited number of forecast periods (perhaps just one) but a large number of small areas. These authors used 1990 population forecasts for thousands of gridcells in San Diego county, California, took random samples of gridcells, amalgamated these sampled gridcells with adjacent cells into zones of certain population sizes, and then calculated the forecast errors for each of the different sets of areas by comparing the forecasts with the 1990 census counts. This enabled predictive distributions of forecast error to be calculated by population size.

Future Directions

In our view there are many directions in which this work can progress. Two main points are mentioned here.

Firstly, more attention could be paid to the way in which migration is handled probabilistically. Sometimes it has been ignored (as in Booth's 2004 forecasts for Australia) or included deterministically in an otherwise probabilistic forecast (e.g. Lee and Tuljapurkar, 1994). Furthermore, net migration rather than gross migration has frequently been forecast, an approach with a number of conceptual and practical shortcomings (Rogers, 1990). Almost no researchers have specifically focused on forecasting gross migration probabilistically, although de Beer (1997) and Wilson and Bell (2004a) are exceptions.

Secondly, further development of probabilistic methods for forecasting ethnic groups, households, the labour force, elderly nursing home requirements, and other demographic variables of relevance to policy-makers, is warranted. The extension of probabilistic methods to specific population groups is important if they are to be adopted by practitioners. National-level population forecasts often provide control totals for more disaggregated forecasts. If, for example, a national statistical office was preparing national and regional population forecasts, it would be unsatisfactory to employ a probabilistic method at the national scale and a deterministic one for the regions. Until these sorts of challenges are addressed, we suspect that the adoption of probabilistic forecasting techniques by statistical offices will remain rather limited.

EXPERIMENTS IN THE PROJECTION OF MIGRATION

Quantitative migration modelling boasts a long-established body of work, particularly in the geographical and regional science journals. Relatively little of this research, however, has been concerned with migration projection, instead focusing on describing and explaining the migration levels and patterns of the recent past. A recent overview of this migration modelling work may be found in Cushing and Poot (2004) and in references given by these authors to reviews of specific migration modelling subfields; here we focus solely on selected methods used in the projection of migration. Internal migration is discussed first, followed by international migration.

Current Practice in Internal Migration Projection

In contrast to many of the models designed to describe and explain past migration patterns, the handling of migration in population projection models is conventionally carried out using simple models – models which exclude non-demographic variables (van Imhoff *et al.*, 1994; Kupiszewski and Kupiszewska, 2003). Typical migration variables include net migration numbers, the origin–destination migration intensities (rates or probabilities) required by multiregional models (Rogers, 1995) or, for the migration pool model, out-migration intensities and in-migration shares (Rees, 1994). Courgeau (1995) proposed an extension to the Rogers matrix model in which the migration rate was defined as the migration flow divided by the product of the origin and destination populations, thus linking the multiregional and gravity modelling traditions. Since their development in the late 1960s and 1970s, multiregional methods have gradually become more and more widespread in the production of subnational population projections. For example, a recent survey of current projection practices in the EU found that a large number of countries now employ the multiregional method, or variants of it, in their official projections (Kupiszewski and Kupiszewska, 2003). However, with migration disaggregated by origin, destination, age and sex, the large and sparse matrices which often arise with these models can render their implementation rather difficult. This is particularly the case with respect to collecting the data, the instability of rates calculated from small numbers, and a large number of variables requiring projection. Strategies for dealing with this issue include: (i) reducing the size of the model by partitioning the migration matrix into several smaller more manageable matrices (Rogers, 1976; Rees, 1997) or by using a biregional or migration pool model; and (ii) using the full origin–destination–age–sex migration matrix but representing it using less information through, for example, parameterised model age schedules of rates (Rogers *et al*., 1978), or the GLIM/IPF approach taken by van Imhoff *et al.* (1997), in which a variety of partial models were tested for their closeness to a full multiregional model. In practice, restrictions on the publication of subnational migration data force this approach upon researchers. In the UK, the most detailed migration data, outside of the one-year interval provided in decennial censuses, consist only of origin–destination, origin–age–sex and destination–age–sex arrays published annually. This matches the optimum model structure suggested by van Imhoff *et al.* (1997).

In most projections, the multiregional transition probabilities are assumed constant, citing their generally stable structure across recent censuses. However, shifts in the generation and distribution of internal migration can have important long-run effects, as pointed out in a comment on time-varying rates in the context of multiregional demography by Plane (1993).

Models that Reduce the Number of Migration Variables Used

As part of the preparation of the 1995-based European Union subnational population projections, van Imhoff *et al.* (1997) set out to discover which parts of the migration matrix were needed to represent accurately the full origin–destination–age–sex migration matrix when 'filled in' by iterative proportional fitting. As such, this is a natural extension of the indirect estimation methods introduced to migration analysis by Willekens (1982) and others (Nair, 1985). Using a log-linear modelling framework, the researchers found that the best trade-off of data requirements to faithful representation of the original full matrix was a model which contained age–sex–origin, age–sex–destination and origin–destination interactions. The researchers also examined the way in which the variables of the best model varied over time, finding that a good fit to the full dataset was provided by time interactions with origin, destination and age–sex only. The important implication from this research is that one can concentrate on relatively few variables in formulating projection assumptions: how national average age–sex interregional migration rates, regional out-migration levels relative to the national level, and regional in-migration shares vary over time. Spatial interaction – the origin–destination effect – was found to be temporally stable. Similar conclusions as to the important time-variant parameters were drawn by Sweeney and Konty (2002) in their analysis and projection of interregional migration in California. The internal migration scenarios prepared for the EU regional projections using the approach designed by van Imhoff *et al.* (1997) are presented in van der Gaag *et al.* (2000).

Projecting Internal Migration Using Predictor Variables

Another strand of migration projection work is the use of regression models which link migration to economic, social and other determinants. Although this line of research is long-established, the accuracy of projected migration from these types of models has often proved disappointing (e.g. Flood *et al*., 1991; van der Gaag *et al*. 2003a, b). In addition, they are frequently very data-hungry, require the assumption that the structural relationships of the past will hold in the future, involve a non-trivial fitting procedure, and require projections of all the independent variables in the regression equation. On the other hand, the advantages of this approach include the linking of projections to theories of migration, and the ability to produce 'what if?' scenarios under different policy assumptions. We mention here one recent application of this approach, a model of interregional migration in the UK designed for 'what if?' policy analysis.

This project to model interregional migration within England and Wales was sponsored by the UK Office of the Deputy Prime Minister, which is responsible for planning and housing policy in England. The results of the study are reported in Champion *et al.* (2002, 2003), Rees *et al.* (2003) and Fotheringham *et al.* (2004). The model operates in three stages. The *first stage* predicts the rate of out-migration using a set of socio-economic characteristics of the origin (informed by a previous review – Champion *et al*., 1998) and calibrated using 15 years of observations pooled for 98 origins. The *second stage* predicts using a smaller set of determinants (but including distance from origin to destination) the distribution of out-migration to destinations. The stage two model is calibrated separately for each origin. The *third stage* of the model takes the stage one parameters and variables and applies them to a set of regional populations to generate predicted out-migration totals. These are then used to predict origin to destination flows using the stage two parameters and variables. Models are fitted at stages one, two and three separately to two sexes and seven broad ages corresponding to life course stages. There is important variation across life course stages in which determinants are significant. Stage three of the model also provides the model user with graphical and cartographic visualisation options. The model was used to test a variety of 'what-if' planning options by matching determinant variable values to the intended effects of policies. Scenarios were developed for 'rescuing Northern conurbations', that is, reducing migration loss from Greater Manchester, Merseyside, West Yorkshire, South Yorkshire and Tyneside, and for 'stimulating the Thames Gateway planning region', building new settlements to house London region migrants. Although the migration model was conceived mainly as a scenario tester, it was demonstrated, in principle, that the migration model could be integrated into a population projection model (Wilson *et al*., 2002).

Projecting International Migration

International migration, particularly at the subnational scale, has generally proven just as difficult to forecast as migration within countries. Indeed, some forms of international migration, such as those between EU member states and between Australia and New Zealand, are more like internal migration flows because of the absence of (legal) migration restrictions. Even countries with policy-determined 'migration programmes' such as Canada and Australia experience large immigration fluctuations from year to year due to the large share of immigration which is unrestricted (such as the immigration of returning residents or nationals). Adding to the challenge is unavailable or poor quality international migration data in countries with otherwise advanced statistical systems. A recent EUROSTAT study has examined ways in which some of these difficulties might be overcome in order to produce better international migration forecasts for EU countries (van der Gaag and van Wissen, 1999; Hilderink *et al*., 2003). These researchers examined numerous relationships between international migration and other variables, including the connections between immigration and various economic indicators, the size of existing migrant subpopulations (as a proxy for migrant networks) and the relationship between asylum applications and immigration. As part of a general framework for international migration forecasting they recommended the modelling of immigration and emigration flows rather than net migration, with immigration classified by motive (employment, asylum) and citizenship, labour migration modelled as a function of GDP per capita, and unemployment rates and emigration classified by citizenship.

Once a model for projecting international flows at national scale has been devised, the question arises of how these flows should be divided between subnational regions. The easiest option is simply to use historical shares. A better option, at least conceptually, is to relate subnational international migration flows to predictor variables. A number of authors have examined this issue for immigration, including Åslund (2005) for Sweden, van der Gaag and van Wissen (2001) for the EU, Hugo (1999) for Australia and Zavodny (1999) for the US. Åslund reports that immigrants are attracted to large cities and those with higher proportions of foreign-born persons. Van der Gaag and van Wissen also found a positive relationship between regional immigration and foreign population stocks, and a negative relationship with the regional unemployment rate. Zavodny's US study at the state scale concluded that the presence of other foreign-born persons is the primary determinant of immigrant location decisions, with economic variables having less influence. Similar work in Europe by Stillwell and colleagues found a close association between current immigration flows and existing stocks and a link with higher internal out-migration (Stillwell *et al*., 1999). Whilst these studies do not provide quantitative models for use in migration projections, they offer analyses which are none the less qualitatively helpful in projection assumption-setting.

PROJECTING DIMENSIONS ADDITIONAL TO AGE, SEX AND REGION

There is considerable interest in the disaggregation of the standard demographic projection, which handles age, sex and region-specific populations, to additional dimensions. Projections of racial or ethnic groups feed into the political debate and internal conflicts within many nation states. This section of the paper reviews contributions to expanding projections to subgroups. It is, of course, simple if the requisite data are available to run parallel projections of non-interacting subgroups. This is frequently done in the projection of racial or ethnic groups (e.g. UK – Immigrant Statistics OPCS, 1979; USA – Hollmann *et al*., 2000, US Census Bureau, 2004; Australia – Taylor, 2003). The number of groups used by the US Census Bureau has expanded considerably with the use of the US Census 2000 population base to include groups of mixed origins.

Projections of Ethnic or Racial Groups

However, researchers have realised that subgroups of the population are not closed populations: they may move from one group to another naturally (e.g. between educational grades); individuals may change their group-identification between surveys or censuses; individuals from different groups may partner and have children of mixed origins; members of different subgroups may share the same living arrangements (i.e. be members of the same household). Some recent progress has been made at introducing these complexities into population projections. Haskey (2000, 2002) has brought together thinking by British demographers to provide advice to the Office for National Statistics (ONS) on a programme of national and local ethnic population projections. The collection of papers (Haskey, 2002) includes proposals for three different projection models. Simpson (2002) proposes a univariate cohort-component model with separate and parallel ethnic projections, as implemented for selected UK local authorities through the POPGROUP software (Andelin and Simpson, 2005). Rees (2002) outlines a multistate model in which new mixed ethnic groups evolve through parenting and childbirth, with an opportunity in adulthood to change groups. Murphy (2002) puts the case for a microsimulation approach in which the demographic drivers of mortality, migration and fertility can be made dependent on ethnic group and other significant socio-economic characteristics. The ONS has investigated the availability of ethnic data for projection and implemented a multiregional cohort-component model, which features ethnic-specific fertility rates, internal migration out-migration intensities, and a distribution model sensitive to ethnic group numbers at destinations (ONS, 2005).

Coleman and Scherbov (2005) have pushed the frontiers of population projection substantially outwards by generating, for the UK, probabilistic projections of the ethnic/foreign origin composition of the population. The projections include the generation of 'complex rapidly growing heterogeneous mixed origin populations' (Coleman and Scherbov, 2005: 1). The UK projections project an expansion of foreign-origin proportions to between 2 and 3 in 10 by mid-century, assuming continuation of current levels of net immigration.

Place of Birth or Origin Projections

Rogers *et al.* (1999) introduced the distinction between foreign-born and native-born into a multiregional projection of the US population. The model expands the multistate model with special rules: birthplace status is fixed from birth, but each birthplace population is subject to a different demographic regime. The authors link their model with a long time series for four US regions of the native and foreign born, and show how the mixes are likely to change in the future. An important feature of the model is the introduction of birthplace-specific migration. Ledent (1980) used US region of birth disaggregations of migration in a multistate projection model, and showed that this additional disaggregation had an impact on the migration forecasts of the model. People born in a region who have moved elsewhere are much more likely to move back there than people born in other regions. Absence of detailed data on migration by place of birth means that this expansion of the multistate projection model has rarely been implemented.

Projections by Educational Status

Recent work has extended population projection into a new dimension, educational status (Lutz *et al*., 1999; Lutz and Goujon, 2001; Goujon and

Lutz, 2004). The idea behind these projections is to forge a link to development. The level and distribution of educational attainment is vital for the future development of any country: in the poorest nations, the gains achieved through making sure that people are literate are very substantial; the richest nations are eager to upgrade the skills of their graduate population through postgraduate training. Lutz and Goujon extend the IIASA world region projections by estimating fertility, mortality and migration for people with completed education and also the expected transitions between educational grades. The results are truly startling: within a few decades China and India will produce more graduates, particularly in technical subjects, than the rest of the world combined (Goujon and Lutz, 2004). This transformation is powering the development of these regions and development in turn stimulates the transformation.

Projections of the Health and Income of the Elderly

Population projections are vital in the articulation of the process of demographic ageing. Demographic ageing is the consequence of past fertility decline and past and future improvement in survival to and through old age. Population ageing can be slowed through heavy immigration/in-migration of younger adults and families to a country or region, and accelerated by rapid emigration/out-migration of younger adults and families. There is a close relationship between future ageing and the income/welfare of the elderly, mediated by the design of pension schemes, welfare benefits and healthcare. Research into the issues associated with ageing has been considerable, so here we pick out a few significant recent contributions. There are two main themes: how much care will a growing elderly population need, and how they will acquire a reasonable income.

The study by a team at the Personal Social Services Research Unit at the University of Kent (Wittenberg *et al*., 2001) uses the national projections for the UK (1998 base forecasts from GAD) as the demographic input. They then disaggregate the old age numbers into dependency groups (needing informal help with domestic tasks, formal community care services, and residential and nursing home care) and forecast costs of providing those services. The results do not make comfortable reading: residential and nursing home places need to expand by 65% over the 1998–2028 period; long-term care spending will rise by 148% in real terms between 1996 and 2031. They assume constant rates of dependency by age, and increases in social care costs of 1% per annum and in healthcare costs of 1.5%. The system will not cope. Several things may happen: dependency rates will lower or shift to older ages and so spread out and reduce demand; or capital (e.g. in housing) will be consumed to pay the additional costs. It is difficult to see how the cost rises can be mitigated as they are a function of the real wage increases in the economy. Work by Egidi (2003) on ageing is more optimistic, at least on the dependency front. She reports on the improving health status of older people across many European countries, although the improvements will be nothing like enough to counteract the 'super-ageing' of the next 30 years as baby-boomers retire.

The income situation of older people in a set of advanced countries was reviewed by Smeeding (2003). Nations adopt a wide range of policies to provide income for older people, and many countries fall well short of the ideal. His expectation, based on a wide range of studies, was that spending on social transfers would fall and that housing, healthcare and long-term care benefits would decrease. This combination of economic and demographic reality makes dismal reading. Guest and McDonald (2000) concurred with this gloomy prognosis for Australia, and Dang *et al.* (2001) for OECD countries, anticipated in an earlier study (OECD, 1996). We need to remind ourselves that the huge projected increases in older people are an indicator of societal success, not failure, and that innovative ways of coping through personal saving, family support and pension reform will be found.

Projecting Specific Labour Force Groups

The size of the available future labour force has been a topic of long-standing interest for policy-makers, and models have been developed both for the population as a whole and for specific occupations. An excellent introduction to these models can be found in Siegel (2002). He included examples of the demographic analysis of the workforces of specific organisations, such as the Armed Forces and the Catholic Church in the US. The Armed Forces face the challenge of recruitment from smaller and smaller youth cohorts in many developed countries, and the Catholic Church faces acute problems of an ageing priesthood and a mismatch between the distributions of churches and members. In many industrialised countries there is anxiety about matching rising demands for healthcare against an ageing workforce. The reports of the UK Medical Workforce Standing Advisory Committee (e.g. Department of Health, 1997) analysed the UK situation, recommending an expansion in the training and retention of doctors and nurses, while at the same time recognising a continuing need to recruit from overseas until the new measures are fully in place (taking perhaps ten years before impacting the supply of labour). A local example of such specific labour force projections and planning was provided by Smith *et al.* (2000), who studied the medical workforce (all grades) of Memphis-Shelby County, Tennessee. They used a scenario-forecasting model to assess the evidence for current and future imbalances between health professional supply and demand.

These examples of projection models for specific workforce groups use for a starting point a basic cohort-component model. Work by researchers at NIDI use multistate models to investigate the interaction of ageing and school teacher supply in the Netherlands (van Imhoff and Henkens, 1998), or to examine the future age structure of the Dutch civil service (Ekamper, 1997). Both models are set within the context of labour market factors – relatively low unemployment but high numbers in early retirement/disability schemes – and government policies – a desire to reduce state expenditure in line with demands for the services delivered by teachers or the civil service and a desire to prolong working life to save on pension spend. By modelling and costing the states of employment (full and part-time), early retirement and disability and the flows between them, the researchers showed the unintended consequences of policy and provided the Dutch Government with a rich source of advice for policy formulation.

Projecting Households

Frequently, population projections are used to generate projections of the numbers and attributes of households, which are small population groups that share living arrangements at the same residential address. The statement made by van Imhoff and Keilman (1991: 4) in their introduction to the LIPRO model still holds true, 14 years on:

> 'Household projection models developed in demography over the past few decades are primarily of the headship rate type . . . Like the labour force participation rate in labour force studies, the headship rate is not a rate in the demographic ("occurrence-exposure") sense and its analytical use reflects a focus on changes in *stocks*, rather than a focus on *flows*.'

The LIPRO ('LIfestyle PROjections') model is based on the methodology of multistate demography, but includes additional features designed to solve the particular problems of household modelling (for an example application, see Murphy and Wang, 1999). It is ironic that, although originally developed for household projections, the LIPRO computer program can be and has been used for a wide range of calculations in multistate demography (Ekamper, 1997; van der Gaag *et al.*, 1997a, b; van Imhoff and Henkens, 1998). There are probably two reasons for the paucity of household applications: lack of agreement about (and social evolution of) the types of households between which people flow, and the lack of suitable data about those flows. Sophisticated register tables or very carefully designed retrospective questions in surveys/censuses are needed.

Studies such as King *et al.* (2000) continue to be the norm. They involve the derivation of the proportions of the population who head particular types of household by age and sex, careful study of the trajectory of key proportions at successive censuses, and the extrapolation of those proportions into the future, subject to appropriate ceilings or floors and consistency checks. A more recent addition to the multistate household modelling toolbox is the ProFamy model developed by Zeng *et al.* (1997, 1999), which differs from LIPRO in terms of the types of transitions which are modelled (McDonald, 2001). Recently Zeng *et al.* (2005) extended this work by addressing the issue of time-varying rates in household modelling. These researchers have proposed a method for preparing time-varying age–sex-specific multistate rates which are consistent with one-sex rates and consistent with projected summary measures of union formation and dissolution.

In addition to the headship rate and multistate methods, there exist other approaches to the preparation of household projections, such as microsimulation. We conclude this section with brief comments on this approach.

Macrosimulation versus Microsimulation Models

Microsimulation is a method that allows for a huge range of population characteristics to be projected and may be regarded as the ultimate level of disaggregation in demographic modelling. Van Imhoff and Post (1998) provided an elegant and comprehensive comparison of macrosimulation models (such as the standard multistate model) and microsimulation models used in population projection. They stress that the models have alternative methods for making similar statements about the future, as the same processes of demographic and socio-economic change are being modelled. We pick out some key points from their summary (van Imhoff and Post, 1998: 132–3):

- In macrosimulation the average outcome is modelled, whereas in microsimulation a sample of outcomes selected randomly are modelled. This distinction may well be disappearing with the development of probabilistic macrosimulation projections.
- Microsimulation models build individual-level determinants of events, while these determinants are external to the model in aggregate macrosimulations.
- Microsimulation gives better results when the state-space is very large, when macrosimulation models become unmanageable because the arrays capturing the interaction between variables become very large and the transition probabilities cannot be estimated reliably.
- Microsimulation can handle interaction effects between variables as the explanatory variables are built in.
- Microsimulation can handle continuous variables and provides richer output.

They observe, however, that microsimulation models are subject to 'specification randomness'. The explanatory variables in the model can take different sets of values depending on the way the die (random number generator) falls when each process is modelled. It is necessary to run the model thousands of times so as to obtain a distribution of outcomes from which median outcomes can be selected. But this is rarely done in practice because of computing time limitations.

In their paper van Imhoff and Post include a review of microsimulation applications in demography up to the late 1990s; more recent examples include the projection of kinship networks in Italy by Tomassini and Wolf (2000) and the DYNAMOD model of the Australian population (Abello *et al*., 2002). A review of selected dynamic microsimulation models (not all of which are primarily demographic) is provided by Zaidi and Rake (2001). Notwithstanding recent studies, the number of applications remains somewhat limited to date (particularly by national statistical offices), due to factors such as the complexity of the modelling, staff resources needed, and substantial data requirements. However, given the current context of an increased availability of micro data-sets (Ballas *et al*., 2005), policy concerns related to population change (particularly ageing) and continued developments in computing power, it would be surprising if microsimulation was not explored further as a tool for making population projections.

SCENARIOS

Whilst probabilistic population forecasts have been shown to be the natural successor to deterministic high, medium and low 'uncertainty' variant projections, scenario projections remain highly valuable tools for understanding population dynamics, and answering 'what if?' questions. A number of examples are discussed.

Migration Scenarios

Probably the most publicised 'what if?' scenarios of recent years were the replacement migration projections of United Nations (2000). This study sought to discover whether higher immigration could be a solution to population decline and ageing in industrialised countries. Five projection scenarios covering the period 1995 to 2050 were computed for eight industrialised countries (France, Germany, Italy, Japan, South Korea, Russia, the UK and the US) and for two global regions, the European Union and Europe. The UN medium variant assumptions were used for fertility and mortality. Amongst the 'what if?' questions asked were: 'What levels of net immigration are needed to maintain the total population at the highest it would reach during the forecast horizon in the absence of migration?' and 'What levels of net immigration are needed to maintain the potential support ratio at its most favourable level reached during the forecast horizon?'. The answer to the first question is that significant additional numbers of migrants are needed except in France and South Korea, and in answer to the second questions: massive and totally implausible levels of migration. The study reports, for example, that total net immigration between 1995–2050 required to maintain the most favourable potential support ratio for South Korea is 5 billion! Although previous work (e.g. Espenshade, 1994) meant that demographers were well aware that migration could prevent population decline, but at plausible levels has little impact on ageing, the strength of the UN study lay in its international comparisons using a consistent method and set of scenario assumptions. It brought attention to the fallacy of using migration to prevent population ageing. Nonetheless, the UN report has attracted much criticism for, amongst other things, giving the impression that immigration is the only solution to ageing in industrialised countries (Coleman, 2002), and its attitude which assumes ageing and population declines are necessarily problems in the first place (Abernethy, 2001; other papers in the March 2001 issue of *Population and Environment* provide further criticisms and discussion).

Whilst the UN study used the same fertility assumptions for each scenario, some researchers have used projection scenarios to investigate the extent to which fertility and migration compensate for one another. In a series of projections for Australia, Kippen and McDonald (1998) assumed a 'medium' assumption of declining mortality and examined various fertility and migration pathways which would lead to stationarity within 100 years. A replacement-level TFR of 2.06 would require no migration; at a TFR of 1.65, annual net immigration would have to be about 100,000, whilst a fertility decline to a TFR of 1.4 would require net immigration to be at about 200,000 per year. Similarly, Lutz and Scherbov (2002) carried out a series of projections for the European Union and reported on the different fertility and migration levels which would give the same total population and elderly dependency ratios (EDR) by 2050. An approximately linear relationship between lower fertility and higher net immigration was found for both total population and the EDR: increasing the TFR by 0.1 was roughly equivalent to increasing the level of net immigration by 100,000 per year.

Bouvier *et al.* (1997) used simulations to highlight the fallacy of assuming that zero net international migration in population projections is equivalent to no international migration. The differences arise due to different age profiles of immigration and emigration. Simulations for the US, Germany and Mexico clearly demonstrated that zero net international migration can influence population growth, both through the direct effects of non-zero net migration for specific ages, and indirectly through births to migrants.

The impact of international migration on the population dynamics and labour force resources of Europe has been explored in a series of scenarios by Bijak *et al.* (2005). The Central European Forum for Migration Research (CEFMR) team project the population of 27 European countries (EU member states except for Cyprus and Malta, plus Norway, Switzerland, Romania and Bulgaria), using a multiregional model, MULTIPOLES, which handles inter-country, intra-European migration using origin to destination transition probabilities and extra-European migration as net immigration numbers. A base scenario includes the most likely international migration developments – increasing migration within Europe and stable migration numbers from outside Europe of 1.2 million people per year. Mortality is assumed to reduce faster than in the UN projections, and a small increase in fertility is expected but is still well below replacement. High and low scenarios are developed which assume higher and lower immigration respectively than the base scenario (combined with the base scenario fertility and mortality assumptions). Four 'what-if' scenarios are also developed: status quo migration (constant current rates or levels), no extra-Europe migration, a projection with immigration numbers needed to maintain a constant population, and three projections with immigration required to maintain constant old-age dependency ratios, a constant economic old-age dependency ratio and a constant labour market dependency ratio. The results (base scenario, low scenario, status quo migration, no extra-Europe migration) show the major challenges of declining numbers of workers and increasing numbers of those retired to social security and care systems. They show that increasing labour force participation (including higher retirement ages) and moderately higher immigration are beneficial. However, the immigration numbers required to reach constant dependency ratios are shown to be completely infeasible (e.g. the annual inflow would need to rise to 36 million per year by mid-century).

Analysis of the Momentum Effect

Projection scenarios have also been used to illuminate the effect of age structure on population change (i.e. population momentum). The effect of population momentum may be isolated by preparing a projection in which mortality rates are held constant, fertility is set to the replacement level throughout the forecast horizon, and there is no migration. Lutz *et al*. (2003a) found that for the European Union (EU15) population momentum recently switched from positive to negative, and will be increasingly negative in the future. Bongaarts and Bulatao (1999) sought to quantify the relative contributions of each of fertility, mortality, migration and age structure in projections for all countries of the world. To achieve this, four sets of projections were produced. After producing a momentum projection, declining mortality was added to give what they termed a *replacement* projection. Then non-replacement fertility was added to give the *natural* projection, before finally incorporating net immigration to produce the *standard* projection. The four projections enabled the contributions of each of the components to be assessed. It was shown, for example, that in the global South nearly half the projected population growth over the course of the twenty-first century will be attributable to momentum.

Fertility Scenarios

O'Neill *et al.* (1999) employed scenario projections to determine the effect of both the length of fertility transition and its eventual long-run level on the long-term population size of current high-fertility countries. They noted that existing long-term projections imply a relatively minor role for the timing of the fertility transition. Taking North Africa as a case study, these researchers demonstrated that the length of fertility transition does indeed play a fairly minor role when the long-run fertility level is high, but is increasingly important the lower the long-run level of fertility.

A similarly insightful research note was provided by Goldstein *et al.* (2003) on the numerical implications of childbearing at increasingly older ages in Europe. They tested the relative importance of two effects. On the one hand, a halt to the upward shift in the ages of childbearing – an end to the tempo effect – would reduce population decline and ageing; on the other hand, with below-replacement fertility, stable population theory suggests slower population decline with longer generational lengths. Through a set of scenario projections they demonstrated that the tempo effect outweighed the longer generation effect, showing that an end to the upward shift in maternity ages slowed population decline and ageing.

CONCLUSIONS

We summarise the conclusions of our survey of population projection developments under the five topics reviewed and then make suggestions about directions that future research might take in this final section.

The activity of **evaluating population forecasts** has increased in recent years. This reflects the growing availability of time series of data to be assessed. Reflection on past errors is a great help in improving methods and assumptions for the future. Analysis of past errors has been used to develop new ways of coping with future uncertainty. However, the purpose of the projection should be borne in mind when comparing errors between projections, as Willekens stressed in his 1990 review. We will learn a great deal from scenario projections, although most will be far from the most likely future.

We have highlighted progress being made in the development of **probabilistic forecasts**, with a healthy current debate on the best methods. There is competition between letting the facts (past time series, past errors) speak for themselves and inputting the (imperfect) knowledge of experts. Past behaviour may not be a good guide to the future in times when trends change. On the other hand, experts may be influenced by other experts and develop a 'herd' view of trends. The history of stockmarkets shows how wrong that can sometimes be.

The development of **migration projections** as part of population forecasting has been relatively neglected, but there is now a rich set of alternative models to be tried out. Getting internal migration and the subnational distribution of international migration reasonably right is vital for accurate subnational forecasts. There are pressures from local, state and regional governments on forecasters to favour their own areas in terms of migration balances, which should be resisted by well justified model designs and assumptions.

A body of research has demonstrated how useful it can be to expand the projection model to track the transitions between categories other than age and region. Important results are now being generated for **educational grades**, **labour force status** and **ethnic groups**. This is bringing to fruition the system of social and demographic statistics envisioned three decades ago by Nobel Laureate Richard Stone (Stone, 1971; United Nations, 1975).

It is also useful to 'think outside the box' about the future. The development of **scenarios** as part of the projection process not only serves to place the preferred or official projections in context, but also helps to demonstrate where particular policies (e.g. encouraging huge immigration to halt Europe's impending population decline) may not work. Scenarios are speculative, but thinking about alternative futures helps to focus activity on the most desirable path.

So, what are the research needs in population projection under the themes covered in this review? We make some suggestions. Greater computational power and the increasing availability of longitudinal micro-level data will help in developing them.

Evaluations. More evaluations of forecast accuracy for *subnational areas* are needed, so that we can continue to improve projection models and assumptions (as in Rees *et al.*, 2001).

Probabilistic projections. The techniques need to be further developed for *subnational forecasts*. There will be particular challenges to overcome: the subnational model will need reduction to fewer parameters whose error distributions can be modelled and used in forecasting. Internal migration is a zero-sum game: the attractiveness of some regions to migrants must be balanced by the non-attractiveness of others. Probabilistic methods also need extension to *group forecasts* (e.g. Coleman and Scherbov, 2005).

Projections of migration. The integration of forecastable explanatory models of *internal migration* into population projections needs to be pursued, so as to provide for scenario forecasting capability. Investigations of period and cohort influences on fertility and mortality are common. Time series are now accumulating which may make such investigation possible for migration. There has been relatively little research into the forecasting of *subnational distribution of international migration* and its links to internal migration, but much research into contemporary migration suggests that the flows are related. Projections will need to deal with increasingly mobile populations. Projections are based for the most part on the usually resident population, located in one home only, but societies also want to know about other populations such as the temporary worker population, visitors, clandestine migrants, the working population or the weekday/weekend population. Statistical agencies will need to develop good measurement instruments for these populations as a precursor to projection.

Adding dimensions to the projection model. There should be more effort devoted to using microsimulation as a method for projecting populations with multiple characteristics, and synergies between macro and micro models will need to be explored.

Scenarios. There is considerable potential with all kinds of projection to adopt a systematic set of scenarios that illuminate what is driving the forecast, and which help assess feasibility of policy.

In conclusion, we predict a bright future for population projection activity because there are so many significant current demographic and social trends which need to be tracked into the future. Good population forecasts are needed as a contribution to debates on issues vital to societies such as population decline, population ageing, low fertility, immigration, changing family structures and living arrangements.

ACKNOWLEDGEMENTS

Financial support for the first-named author for this paper was provided in part by a collaborative research agreement with the Queensland Government's Office of Economic and Statistical Research.

REFERENCES

Abello A, Kelly S, King A. 2002. *Demographic Projections with DYNAMOD-2*. NATSEM Technical Paper No. 21. National Centre for Social and Economic Modelling, University of Canberra. http://www.natsem.canberra.edu.au/publication.jsp?titleID=TP21 [accessed 11 June 2005].

Abernethy VD. 2001. Comment on Bermingham's summary of the UN's year 2000 'Replacement Migration: Is it a Solution to Declining Population and Aging?' *Population and Environment* **22**(4): 365–375. DOI: 10.1023/A:1006741420885

Alders M, de Beer J. 2004. Assumptions on fertility in stochastic population forecasts. *International Statistical Review* **72**: 65–79.

Alho JM. 1997. Scenarios, uncertainty and conditional forecasts of the world population. *Journal of the Royal Statistical Society Series A* **160**: 71–85.

Alho JM. 2002a. *The population of Finland in 2050 and beyond*. Discussion paper no. 826. Research Institute of the Finnish Economy: Helsinki. http://www.etla.fi/files/648_dp826.pdf [accessed 14 June 2005].

Alho J. 2002b. Stochastic forecast of the Lithuanian population 2001–2050. In *The Economic Effects of Ageing and Demographic Uncertainty in Lithuania*, Alho JM, Jensen SHE, Lassila J, Lazutka R, Morkuiene A, Valkonen T. Research Institute of the Finnish Economy: Helsinki. http://www.etla.fi/lithuania/ [accessed 14 June 2005].

Alho JM, Spencer BD. 1997. The practical specification of the expected error of population forecasts. *Journal of Official Statistics* **13**: 203–225. http://www.jos.nu/ [accessed 14 June 2005].

Alho JM, Nikander T. 2004. *Uncertain Population of Europe-Summary Results from a Stochastic Forecast*. Deliverable 12, Work Package 6, EU 5th Framework funded research project Uncertain Population of Europe (UPE), HPSE-CT-2001-00095. http://www.stat.fi/tup/euupe/rp_reports_and_pub.html [accessed 12 June 2005].

Andelin J, Simpson S. 2005. POPGROUP Population Estimation and Forecasting System Reference Manual. Cathie Marsh Centre for Census and Survey Research: Manchester.

Åslund O. 2005. Now and forever? Initial and subsequent location choices of immigrants. *Regional Science and Urban Economics* **35**: 141–165. DOI: 10.1016/j.regsciurbeco.2004.02.001

Ballas D, Rossiter D, Thomas B, Clarke G, Dorling D. 2005. *Geography Matters: Simulating the Local Impacts of National Social Policies*. Joseph Rowntree Foundation: York. http://www.jrf.org.uk/bookshop/eBooks/1859352669.pdf [accessed 14 June 2005].

Bijak J, Kupiszewska D, Kupiszewski M, Saczuk K. 2005. Impact of international migration on population dynamics and labour force resources in Europe. CEFMR Working Paper 1/2005. Central European Forum for Migration Research: Warsaw. http://www.cefmr.pan.pl/docs/cefmr_wp_2005-01.pdf [accessed 11 June 2005].

Bongaarts J, Bulatao RA. 1999. Completing the demographic transition. *Population and Development Review* **25**: 515–529. DOI: 10.1111/j.1728-4457.1999.00515.x

Booth H, Maindonald J, Smith L. 2002. Applying Lee-Carter under conditions of variable mortality decline. *Population Studies* **56**: 325–336. DOI: 10.1080/003247202762664900

Booth H. 2004. On the importance of being uncertain: forecasting population futures for Australia. *People and Place* **12**(2): 1–12. http://elecpress.monash.edu.au/pnp/view/issue/?volume=12&issue=2 [accessed 14 June 2005].

Bouvier LF, Poston DL, Zhai NB. 1997. Population growth impacts of zero net international migration. *International Migration Review* **31**: 294–311.

Champion T, Fotheringham AS, Boyle P, Rees P, Stillwell J. 1998. *The Determinants of Migration Flows in England: A Review of Existing Data and Evidence*. Report prepared for the Department of the Environment, Transport and the Regions. http://www.geog.leeds.ac.uk/publications/DeterminantsOfMigration/report.pdf [accessed 14 June 2005].

Champion T, Fotheringham S, Rees P, Bramley G, *et al.* 2002. *Development of a Migration Model*. Office of the Deputy Prime Minister: London. http://www.odpm.gov.uk/stellent/groups/odpm_housing/documents/page/odpm_house_601865.pdf [accessed 11 June 2005].

Champion T, Bramley G, Fotheringham S, Macgill J, Rees P. 2003. A migration modelling system to support government decision-making. In *Planning Support Systems in Practice*, Geertman S, Stillwell J (eds). Springer: Berlin; 269–290.

Coleman DA. 2002. Replacement migration, or why everyone is going to have to live in Korea: a fable for our times from the United Nations. *Philosophical Transactions of the Royal Society of London B* **357**: 583–598.

Coleman DA, Scherbov S. 2005. Immigration and ethnic change in low-fertility countries – towards a new demographic transition. Paper presented at the

Population Association of America Annual Meeting, Philadelphia, March 31–April 2. 2005, Session 98, 1 April. http://www.apsoc.ox.ac.uk/Oxpop/publications%20files/WP29.pdf [accessed 14 June 2005].

Courgeau D. 1995. Migration theories and behavioural models. *International Journal of Population Geography* **1**: 19–27.

Cruijsen H, Keilman N. 1984. *Prognose van de bevolking van Nederland na 1980. Deel 2: Modelbouw en hypothesevorming*. Staatsuitgeverij: The Hague.

Cushing B, Poot J. 2004. Crossing boundaries and borders: regional science advances in migration modelling. *Papers in Regional Science* **83**: 317–338. DOI: 10.1007/s10110-003-0188-5

Dang TT, Antolin P, Oxley H. 2001. Fiscal implications of ageing: projections of age-related spending. *OECD Economics Department Working Paper 305*. OECD: Paris. http://www.olis.oecd.org/olis/2001doc.nsf/linkto/eco-wkp(2001)31. Also published in *OECD Economic Outlook* June 2001: **1**. IV. Fiscal Implications of Ageing: Projections of Age-Related Spending. http://www.oecd.org/dataoecd/1/22/2085481.pdf [accessed 14 June 2005].

de Beer J. 1997. The effect of uncertainty of migration on national population forecasts: the case of the Netherlands. *Journal of Official Statistics* **13**: 227–243. http://www.jos.nu/ [accessed 14 June 2005].

de Beer J. 2000. *Dealing with uncertainty in population forecasting*. Demographic Working Paper, Statistics Netherlands. http://www.cbs.nl/en/publications/articles/population-society/population/dealing-with-uncertainty.pdf [accessed 14 June 2005].

de Beer J, Alders M. 1999. *Probabilistic population and household forecasts for the Netherlands*. Paper prepared for the joint ECE-Eurostat work session on demographic projections, Perugia, Italy, 3–7 May 1999. http://www.unece.org/stats/documents/1999/05/projections/45.e.pdf [accessed 14 June 2005].

Department of Health. 1997. *Planning the Medical Workforce*. Medical Workforce Standing Advisory Committee: Third Report. http://www.dh.gov.uk/assetRoot/04/03/91/27/04039127.pdf [accessed 14 June 2005].

Egidi V. 2003. Health status of older people. *Genus* **LIX**(1): 169–200.

Ekamper P. 1997. Future age-conscious manpower planning in the Netherlands. *International Journal of Manpower* **18**: 232–247.

Espenshade TJ. 1994. Can immigration slow US population aging? *Journal of Policy Analysis and Management* **13**: 759–768.

Flood J, Maher C, Newton P, Roy J. 1991. *The Determinants of Internal Migration in Australia*. CSIRO: Melbourne.

Fotheringham AS, Rees P, Champion T, Kalogirou S, Tremayne AR. 2004. The development of a migration model for England and Wales I: overview and modelling out-migration. *Environment and Planning A* **36**: 1633–1672. DOI: 10.1068/a36136

GAD. 2004. Projections Database. http://www.gad.gov.uk/Population/index.asp [accessed 11 June 2005].

Gans de HA. 1999. *Population Forecasting 1895–1945: The Transition to Modernity*. Kluwer Academic: Dordrecht.

Goldstein J, Lutz W, Scherbov S. 2003. Long-term population decline in Europe: the relative importance of tempo effects and generational length. *Population and Development Review* **29**: 699–707. DOI: 10.1111/j.1728-4457.2003.00699.x

Goujon A, Lutz W. 2004. Future human capital: Population projections by level of education. In *The End of World Population Growth in the 21st Century: New Challenges for Human Capital Formation and Sustainable Development*, Lutz W, Sanderson WC, Scherbov S (eds). Earthscan: London; 121–157.

Guest RS, McDonald IM. 2000. Population ageing and projections of government social outlays in Australia. *Australian Economic Review* **33**: 49–64. DOI: 10.1111/1467-8462.00135

Gullickson A. 2001. *Multiregional probabilistic forecasting*. Paper presented at the Young Scientists Summer Program Midsummer Workshop, International Institute for Applied Systems Analysis, July 2001, Vienna, Austria.

Gullickson A, Moen J. 2001. *The use of stochastic methods in local area population forecasts*. Paper presented at the annual meeting of the Population Association of America, March 2001, Washington, DC.

Haskey J. 2000. Projections of the population by ethnic group: a sufficiently interesting or definitely necessary exercise to undertake? *Population Trends* **102**: 34–40. http://www.statistics.gov.uk/downloads/theme_population/PT102book_v3.pdf [accessed 14 June 2005].

Haskey J (ed.). 2002. *Population Projections by Ethnic Group: A Feasibility Study*. Stationery Office: London. http://www.statistics.gov.uk/downloads/theme_population/SMPS_67_v2.pdf [accessed 11 June 2005].

Heering L, Keilman N. 2004. *Expert Views on Population Forecasts and their Uncertainty*. Report for Project *Stochastic Population Projections for the EEA*, Eurostat. Available from the first author at Heering@nidi.nl.

Hilderink H, van der Gaag N, van Wissen L, Jennissen R, Roman A, Salt J, Clarke J, Pinkerton C. 2003. Analysis and forecasting of international migration by major groups (part III). *Eurostat Working Paper 3/2002/E/no.17*, European Commission, Luxembourg. http://epp.eurostat.cec.eu.int/cache/ITY_

OFFPUB/KS-CC-03-002/EN/KS-CC-03-002-EN.PDF [accessed 14 June 2005].

Hollmann FW, Mulder TJ, Kallan JE. 2000. *Methodology and Assumptions for the Population Projections of the United States: 1999 to 2100*. Population Division Working Paper No. 38, US Census Bureau, Department of Commerce, Washington, DC. http://www.census.gov/population/www/documentation/twps0038.pdf [accessed 14 June 2005].

Hugo G. 1999. Regional development through immigration? The reality behind the rhetoric. *Research Paper 9 1999–2000*, Parliament of Australia, Canberra. http://www.aph.gov.au/library/pubs/rp/1999-2000/2000rp09.htm [accessed 14 June 2005].

Immigrant Statistics OPCS. 1979. Population of New Commonwealth and Pakistani ethnic origin: new projections. *Population Trends* **16**: 22–27.

Keilman N. 1997. Ex-post errors in official population forecasts in industrialized countries. *Journal of Official Statistics* **13**: 245–277. http://www.jos.nu/ [accessed 14 June 2005].

Keilman N. 1999. How accurate are the United Nations world population projections? In *Frontiers of Population Forecasting*, Lutz W, Vaupel JW, Ahlburg DA (eds). Population Council: New York; 15–41. Reprinted supplement to *Population and Development Review* **24**(4).

Keilman N. 2001. Data quality and accuracy of United Nations population projections, 1950–95. *Population Studies* **55**: 149–164.

Keilman N, Pham DQ. 2000. Predictive intervals for age-specific fertility. *European Journal of Population* **16**: 41–66.

Keilman N, Pham DQ. 2004. *Empirical errors and predicted errors in fertility, mortality and migration forecasts in the European Economic Area.* Discussion Papers 386. Statistics Norway: Oslo. http://www.ssb.no/publikasjoner/DP/pdf/dp386.pdf [accessed 12 June 2005].

Keilman N, Pham DQ, Hetland A. 2001. *Norway's uncertain demographic future.* Statistics Norway Social and Economic Studies no. 105. Statistics Norway: Oslo. http://www.ssb.no/english/subjects/02/03/sos105_en/ [accessed 14 June 2005].

King D, Hayden J, Jackson R, Holmans A, Anderson D. 2000. Population of households in England to 2021. *Population Trends* **99**: 13–19. http://www.statistics.gov.uk/downloads/theme_population/PT99book.pdf [accessed 14 June 2005].

Kippen R, McDonald P. 1998. Achieving population targets for Australia: an analysis of the options. *People and Place* **6**(2): 11–23. http://elecpress.monash.edu.au/pnp/free/pnpv6n2/mcdkipp3.htm [accessed 14 June 2005].

Kupiszewski M, Kupizewska D. 2003. *Internal migration component in subnational population projections in member states of the European Union.* Working Paper 2/2003, Central European Forum for Migration Research, Warsaw. http://www.cefmr.pan.pl/docs/cefmr_wp_2003-02.pdf [accessed 11 June 2005].

Ledent J. 1980. *Constructing multiregional life tables using place-of-birth-specific migration data.* Working Paper WP-80-096, International Institute for Applied Systems Analysis. http://www.iiasa.ac.at/cgi-bin/pubsrch?WP80096 [accessed 14 June 2005].

Lee R. 1993. Modeling and forecasting the time series of US fertility: age patterns, range and ultimate level. *International Journal of Forecasting* **9**: 187–202. DOI: 10.1016/0169-2070(93)90004-7

Lee R. 1999. Probabilistic approaches to population forecasting. In *Frontiers of Population Forecasting*, Lutz W, Vaupel JW, Ahlburg DA (eds). Population Council: New York; 156–190. Reprinted supplement to *Population and Development Review* **24**(4).

Lee R. 2000. The Lee-Carter method for forecasting mortality, with various extensions and applications. *North American Actuarial Journal* **4**: 80–93.

Lee RD, Carter LR. 1992. Modeling and forecasting US mortality. *Journal of the American Statistical Association* **87**: 659–671.

Lee RD, Tuljapurkar S. 1994. Stochastic population forecasts for the United States: beyond high, medium and low. *Journal of the American Statistical Association* **89**: 1175–1189.

Lee R, Tuljapurkar S. 1998. Uncertain demographic futures and social security finances. *American Economic Review* **88**: 237–241.

Lee R, Miller T, Edwards RD. 2003. *The growth and aging of California's population: demographic and fiscal projections, characteristics and service needs.* Special Report, California Policy Research Center, University of California, Berkeley, US. http://www.ucop.edu/cprc/pubsspser.html#sb910 [accessed 14 June 2005].

Leiwen J, O'Neill BC. 2004. Toward a new model for probabilistic household forecasts. *International Statistical Review* **72**: 51–64.

Leslie PH. 1945. On the use of matrices in certain population dynamics. *Biometrika* **33**: 183–212.

Leslie PH. 1948. Some further notes on the use of matrices in population mathematics. *Biometrika* **35**: 213–245.

Lutz W, Sanderson W, Scherbov S. 1996. Probabilistic population projections based on expert opinion. In *The Future Population of the World: What Can We Assume Today?* (2nd edn), Lutz W (ed.). Earthscan: London; 397–428.

Lutz W, Sanderson W, Scherbov S. 1997. Doubling of world population unlikely. Nature **387**: 803–805.

Lutz W, Scherbov S. 1998. An expert-based framework for probabilistic national population projections: the example of Austria. *European Journal of Population* **14**: 1–17.

Lutz W, Scherbov S. 1999. First probabilistic population projections for the European Union. In *Compendium of Family Studies in Austria 1999*, Lutz W (eds). Austrian Institute for Family Studies: Vienna; 123–139.

Lutz W, Goujon A, Doblhammer-Reiter G. 1999. Demographic dimensions in forecasting: Adding education to age and sex. In *Frontiers of Population Forecasting*, Lutz W, Vaupel J, Ahlburg D (eds). The Population Council: New York; 42–58. Supplement to *Population and Development Review* **24**(4).

Lutz W, Goujon A. 2001. The world's changing human capital stock: Multi-state population projections by educational attainment. *Population and Development Review* **27**: 323–339. DOI: 10.1111/j.1728-4457.2001.00323.x

Lutz W, Sanderson W, Scherbov S. 2001a. The end of world population growth. *Nature* **412**(6846): 543–545. DOI: 10.1038/35087589

Lutz W, Scherbov S, van Aardt C. 2001b. South Africa's uncertain demographic present and future. *Interim Report IR-01-019*, International Institute for Applied Systems Analysis, Laxenburg, Austria. http://www.iiasa.ac.at/Publications/Documents/IR-01-019.pdf. [Accessed 18 June 2005].

Lutz W, Scherbov S. 2002. Can immigration compensate for Europe's low fertility? *Interim Report IR-02-052*, International Institute for Applied Systems Analysis, Laxenburg, Austria. http://www.iiasa.ac.at/Publications/Documents/IR-02-052.pdf.

Lutz W, O'Neill BC, Scherbov S. 2003a. Europe's population at a turning point. *Science* **299**: 1991–1992. DOI: 10.1126/science.1080316

Lutz W, Scherbov S, Sanderson W. 2003b. The end of population growth in Asia. *Journal of Population Research* **20**: 125–141.

Lutz W, Scherbov S. 2004a. Probabilistic population projections for Singapore and Asia. *Innovation* **5**(1): 44–45.

Lutz W, Scherbov S. 2004b. Probabilistic population projections for India with explicit consideration of the education-fertility link. *International Statistical Review* **72**: 81–92.

Lutz W, Sanderson WC, Scherbov S (eds). 2004. *The End of World Population Growth in the 21st Century.* Earthscan: London.

Mathematical Population Studies 5.3. 1995. Special issue on *Population projections: simple versus complex models.*

McDonald P. 2001. *Medium and Long-term Projections of Housing Needs in Australia: Positioning Paper.* Australian Housing and Urban Research Institute, Canberra. http://www.ahuri.edu.au/global/docs/pp_projections.pdf?CFID=215249&CFTOKEN=38959577

Miller T. 2002. California's uncertain population future. Unpublished paper, Department of Demography, University of California, Berkeley, US. Appendix to Lee, Miller and Edwards (2003). http://www.demog.berkeley.edu/~tmiller/papers/p2002.california.forecast.pdf [Accessed 7 June 2004].

Murphy M. 2002. Consistency in considerations for population projections by ethnic group, and the possible role of microsimulation models. In *Population Projections by Ethnic Group: A Feasibility Study*, Haskey J (ed.). Stationery Office: London; 17–26. http://www.statistics.gov.uk/downloads/theme_population/SMPS_67_v2.pdf [Accessed 14 June 2005].

Murphy M, Wang D. 1999. Forecasting British families into the twenty-first century. In *Changing Britain: Families and Households in the 1990s*, McRae S (ed.). Oxford University Press: Oxford; 100–137.

Nair PS. 1985. Estimation of period-specific gross flows from limited data. *Demography* **22**: 133–142.

National Research Council (Panel on Population Projections). 2000. *Beyond Six Billion: Forecasting the World's Population*. National Academy Press: Washington, DC.

O'Neill BC. 2004. Conditional probabilistic population projections: an application to climate change. *International Statistical Review***72**: 167–184.

O'Neill BC, Scherbov S, Lutz W. 1999. The long-term effect of the timing of fertility decline on population size. *Population and Development Review* **25**: 749–756. DOI: 10.1111/j.1728-4457.1999.00749.x

ONS. 2005. *Ethnic Projections. Proposed Methodology: Summary Report.* Office for National Statistics, Titchfield. Available from Pete Large, Ethnic Projections Project, e-mail Pete.Large@ons.gsi.gov.uk.

OECD. 1996. *Ageing in OECD Countries: A Critical Policy Challenge.* Social Policy Studies no.20. Organisation for Economic Development and Co-operation: Paris.

Plane DA. 1993. Requiem for the fixed-transition-probability migrant. *Geographical Analysis* **25**: 211–223.

Population and Environment **22**(4), March 2001, 355–409. DOI: 10.1023/A:1006782904046 to DOI: 10.1023/A:1006749722702

Rees P. 1994. *Key issues in subnational projection models.* Working Paper 94/8, School of Geography, University of Leeds.

Rees P. 1997. Problems and solutions in forecasting geographical populations. *Journal of the Australian Population Association* **14**: 145–166.

Rees P. 2002. New models for projecting UK ethnic group populations at national and subnational scales. In *Population Projections by Ethnic Group: A Feasibility Study*, Haskey J (ed.). Stationery Office:

London: 27–51. http://www.statistics.gov.uk/downloads/theme_population/SMPS_67_v2.pdf [Accessed 14 June 2005].

Rees P, Turton I. 1998. Investigation of the effects of input uncertainty on population forecasting. Paper prepared for the GeoComputation 98 Conference, Bristol, UK, 17–19 September 1998. http://www.geocomputation.org/1998/23/gc_23.htm [accessed 18 June 2005].

Rees P, Kupiszewski M, Eyre H, Wilson T, Durham H. 2001. *The evaluation of regional population projections for the European Union.* Eurostat Working Paper, Population and Social Conditions 3/2001/E/no 9. European Commission: Luxembourg. Catalogue No. KA-AP-01-041-EN-I. http://epp.eurostat.cec.eu.int/portal/page [Accessed 18 June 2005].

Rees P, Fotheringham AS, Champion T. 2003. Modelling migration for policy analysis. In *Applied GIS and Spatial Analysis*, Stillwell J, Clarke G (eds). John Wiley & Sons: Chichester: 259–296.

Renshaw AE, Haberman S. 2003. Lee-Carter mortality forecasting with age-specific enhancement. *Insurance: Mathematics and Economics* **33**: 255–272. DOI: 10.1016/S0167-6687(03)00138-0

Rogers A. 1966. Matrix methods of population analysis. *Journal of the American Institute of Planners* **32**(1): 40–44.

Rogers A. 1975. *Introduction to Multiregional Mathematical Demography.* John Wiley: New York.

Rogers A. 1976. Shrinking large-scale population projection models by aggregation and decomposition. *Environment and Planning A* **8**: 515–541.

Rogers A. 1986. Parameterised multistate population dynamics and projections. *Journal of the American Statistical Association* **81**: 48–61.

Rogers A. 1990. Requiem for the net migrant. *Geographical Analysis* **22**: 283–300.

Rogers A. 1995. *Multiregional Demography: Principles, Methods and Extensions.* John Wiley: Chichester.

Rogers A, Raquillet R, Castro LJ. 1978. Model migration schedules and their applications. *Environment and Planning A* **10**: 475–502.

Rogers A, Willekens F. 1986. *Migration and Settlement: A Multiregional Comparative Study.* Reidel: Dordrecht.

Rogers A, Little J, Raymer J. 1999. Disaggregating the historical sources of regional foreign-born and native-born population growth in the United States: a new method with applications. *International Journal of Population Geography* **5**: 449–475.

Sanderson WC, Scherbov S, Lutz W, O'Neill BC. 2004. Applications of probabilistic population forecasting. In *The End of World Population Growth in the 21st Century*, Lutz W, Sanderson WC, Scherbov S (eds). Earthscan: London; 85–119.

Schoen R. 1988. *Modeling Multigroup Populations.* Plenum Press: New York.

Shaw C. 2004. Interim 2003-based national population projections for the United Kingdom and constituent countries. *Population Trends* **118**: 6–16. http://www.statistics.gov.uk/downloads/theme_population/PT118_V1.pdf [Accessed 18 June 2005].

Siegel JS. 2002. The demography of the labor force and the workforce of organisations. In *Applied Demography: Applications to Business, Government, Law and Public Policy.* Academic Press: San Diego; 329–395.

Simpson L. 2002. Estimating methodologies and assumptions for projections: current practice for population projections by ethnic group. In *Population Projections by Ethnic Group: A Feasibility Study*, Haskey J (ed.). Stationery Office: London; 53–72. http://www.statistics.gov.uk/downloads/theme_population/SMPS_67_v2.pdf [Accessed 18 June 2005].

Simpson S (ed.). 1998. *Making Local Population Statistics: A Guide for Practitioners.* Local Authorities Research and Intelligence Association: Wokingham. Available from LARIA, 9 Cortland Road, Nunthorpe, Middlesbrough, TS7 OJX. E-mail: lariaoffice@aol.com.

Smeeding T. 2003. Income maintenance in old age: current status and future prospects for rich countries. *Genus* **LIX**(1): 51–83.

Smith KD, Perez-Johnson I, Wooldridge J. 2000. Uncertainty and forecasting local health professional shortages. *Population Research and Policy Review* **19**: 477–503. DOI: 10.1023/A:1010700930344

Smith SK, Sincich T. 1988. Stability over time in the distribution of population forecast errors. *Demography* **25**: 461–474.

Smith SK, Shahidullah M. 1995. An evaluation of population projection errors for census tracts. *Journal of the American Statistical Association* **90**: 64–71.

Smith SK, Tayman J, Swanson DA. 2001. *State and Local Population Projections: Methodology and Analysis.* Kluwer Academic: New York.

Smith SK, Tayman J. 2003. An evaluation of population projections by age. *Demography* **40**: 741–757.

Smith SK, Tayman J. 2004. Confidence intervals for population forecasts: a case study of time series models for states. Paper prepared for the annual meeting of the Population Association of America, Boston, 1–3 April 2004. http://paa2004.princeton.edu/download.asp?submissionId=40269 [Accessed 19 June 2005].

Statistics Finland. 2004. *A Forecast of the Population of Europe 2004–2050.* http://www.stat.fi/tup/euupe/ [Accessed 12 June 2005].

Stillwell J, Eyre H, Rees PH, Bucher H, Salt J, Clarke J. 1999. *Regional International Migration and Inter-Regional Migration.* Final Report prepared for the European Commission Directorate General XVI (Regional Policy and Cohesion) and EUROSTAT

(Statistical Office of the European Communities) for EDRF Study No. 98/00/27/174.

Stone R. 1971. An integrated system of demographic manpower and social statistics and its links with the system of national economic accounts. *Sankhya B* **33**: 1–184.

Stoto MA. 1983. The accuracy of population projections. *Journal of the American Statistical Association* **78**: 13–20.

Sweeney SH, Konty KJ. 2002. Population forecasting with nonstationary multiregional growth matrices. *Geographical Analysis* **34**: 289–312.

Taylor J. 2003. Population futures in the Australian desert. *Australian Geographer* **34**: 355–370. DOI: 10.1080/0004918032000152438

Tayman J, Swanson DA. 1996. On the utility of population forecasts. *Demography* **33**: 523–528.

Tayman J, Schafer E, Carter L. 1998. The role of population size in the determination and prediction of population forecast errors: an evaluation using confidence intervals for subcounty areas. *Population Research and Policy Review* **17**: 1–20. DOI: 10.1023/A:1005766424443

Tayman J, Swanson DA. 1999. On the validity of MAPE as a measure of population forecast accuracy. *Population Research and Policy Review* **18**(4): 299–322. DOI: 10.1023/A:1006166418051

Tayman J, Swanson DA, Barr CF. 1999. In search of the ideal measure of accuracy for subnational demographic forecasts. *Population Research and Policy Review* **18**(5): 387–409. DOI: 10.1023/A: 1006317430570

Tomassini C, Wolf DA. 2000. Shrinking kin networks in Italy due to sustained low fertility. *European Journal of Population* **16**: 353–372.

Tuljapurkar S, Li N, Boe C. 2000. A universal pattern of mortality decline in the G7 countries. *Nature* **405**(6788): 789–792. DOI: 10.1038/35015561

United Nations. 1975. *Towards a System of Social and Demographic Statistics.* Studies in Methods, Series F, no. 18. Document ST/EAS/STAT/ser.F18. Department of Economic and Social Affairs, Statistical Office. United Nations: New York.

United Nations. 2000. *Replacement Migration: Is it a Solution to Declining and Ageing Populations?* Publication ESA/P/WP.160. Population Division, Department of Economic and Social Affairs. United Nations: New York. http://www.un.org/esa/population/publications/migration/migration.htm [Accessed 19 June 2005].

United Nations. 2004. *World Population Prospects: The 2002 Revision. Volume III: Analytical Report.* United Nations: New York. Executive summary available via http://www.un.org/esa/population/publications/wpp2002/WPP2002_Vol3.htm [Accessed 19 June 2005].

US Census Bureau. 2004. *US Interim Projections by Age, Sex, Race, and Hispanic Origin.* http://www.census.gov/ipc/www/usinterimproj/ [Accessed 19 June 2005].

van der Gaag N, van Imhoff E, van Wissen L. 1997a. *Internal Migration in the Countries of the European Union. Report with Full Tables.* Netherlands Interdisciplinary Demographic Institute: The Hague.

van der Gaag N, van Imhoff E, van Wissen L. 1997b. *Long-term Internal Migration Scenarios for the Countries of the European Union.* Eurostat Working Paper 5/1997/E/no.5. European Commission: Luxembourg.

van der Gaag N, van Wissen L. 1999. *Analysis and Forecasting of International Migration by Major Groups (Part II).* Eurostat Working Paper 3/1999/E/no.9. European Commission: Luxembourg.

van der Gaag N, van Imhoff E, van Wissen L. 2000. Internal migration scenarios and regional population projections for the European Union. *International Journal of Population Geography* **6**: 1–19. DOI: 10.1002/(SICI)1099-1220(200001/02)6:1 < 1::AID-IJPG170>3.0.CO;2-3

van der Gaag N, van Wissen L. 2001. Determinants of the subnational distribution of immigration. *Tijdschrift voor economische en sociale geografie* **92**: 27–41.

van der Gaag N, van Wissen L, Rees P, Stillwell J, Kupiszewski M. 2003a. *Study of Past and Future Interregional Migration Trends and Patterns within European Union countries: In Search for a Generally Applicable Explanatory Model.* Report to Eurostat, Netherlands Interdisciplinary Demographic Institute: The Hague, Netherlands.

van der Gaag N, van Wissen L, Stillwell J, Rees P, Kupiszewski M. 2003b. *In search for a general model of interregional migration.* Paper prepared for the European Population Conference, 26–30 August 2003, Warsaw, Poland.

van Imhoff E. 1990. The exponential multidimensional demographic projection model. *Mathematical Population Studies* **2**(3): 171–182.

van Imhoff E, Keilman NW. 1991. *LIPRO 2.0: An Application of a Dynamic Demographic Projection Model to Household Structure in The Netherlands.* NIDI CBGS Publications no. 23. Swets & Zeitlinger: Amsterdam/Lisse. http://www.nidi.knaw.nl/en/output/nidicbgs/nidicbgs-publ-23.pdf/nidicbgs-publ-23.pdf [Accessed 14 June 2005].

van Imhoff E, Henkens K. 1998. The budgetary dilemmas of an ageing workforce: a scenario study of the public sector in the Netherlands. *European Journal of Population* **14**: 39–59. http://www.nidi.knaw.nl/en/output/1998/ejp-14-01-vanimhoff.pdf/ejp-14-01-vanimhoff.pdf [Accessed 14 June 2005].

van Imhoff E, Post W. 1998. Microsimulation methods for population projection. *Population: An English Selection* **10**: 97–138.

van Imhoff E, van Wissen L, Spiess K. 1994. *Regional Population Projections in the Countries of the European Economic Area*. NIDI/CBGS Publications no. 31, Swets & Zeitlinger: Amsterdam/Lisse.

van Imhoff E, van der Gaag N, van Wissen L, Rees P. 1997. The selection of internal migration models for European regions. *International Journal of Population Geography* **3**: 137–159. DOI: 10.1002/(SICI)1099-1220(199706)3:2<137::AID-IJPG63>3.0.CO;2-R

Whelpton PK. 1936. An empirical method for calculating future population. *Journal of the American Statistical Association* **31**: 457–473.

Willekens F. 1982. Multidimensional population analysis with incomplete data. In *Multidimensional Mathematical Demography*, Land KC, Rogers A (eds). Academic Press: New York; 43–111.

Willekens FJ. 1990. Demographic forecasting: state-of-the-art and research needs. In *Emerging Issues in Demographic Research*, Hazeu CA, Frinking CAB (eds). Elsevier Science: Amsterdam; 9–66.

Wilmoth JR. 1996. Mortality projection in Japan: a comparison of four methods. In *Health and Mortality Among Elderly Populations*, Caselli G, Lopez AD (eds). Clarendon Press: Oxford; 266–287.

Wilson T. 2004. *The application of a probabilistic framework to New Zealand's official national population projections*. QCPR Discussion Paper 2004/02, School of Geography, Planning and Architecture, University of Queensland, Brisbane, Australia. http://www.geosp.uq.edu.au/qcpr/Homepage/discussion_papers/2004-02_FrameworkNZ.pdf [Accessed 11 June 2005].

Wilson T, Bell M. 2004a. Australia's uncertain demographic future. *Demographic Research* **11**, article 8. http://www.demographic-research.org/ [Accessed 19 June 2005].

Wilson T, Bell M. 2004b. *Probabilistic regional population forecasts: the example of Queensland, Australia*. QCPR Discussion Paper 2004/04, Queensland Centre for Population Research, School of Geography, Planning and Architecture, University of Queensland, Brisbane, Australia. http://www.geosp.uq.edu.au/qcpr/Homepage/discussion_papers/2004-04_ProbabilisticRegionPopulationForecast.pdf [Accessed 11 June 2005].

Wilson T, Hollis J, Rees P. 2002. Providing a population projection framework. In *Development of a Migration Model*, Champion T, Fotheringham S, Rees P, Bramley G, *et al*. Office of the Deputy Prime Minister: London; 239–249. http://www.odpm.gov.uk/stellent/groups/odpm_housing/documents/page/odpm_house_601865.pdf [Accessed 11 June 2005].

Wittenberg R, Pickard L, Comas-Herrera A, Davies B. 2001. Demand for long-term care for older people in England to 2031. *Health Statistics Quarterly* **12**: 5–17. http://www.statistics.gov.uk/downloads/theme_health/HSQ12_v2.pdf [Accessed 19 June 2005].

Wolf DA. 2004. Another variation on the Lee-Carter model. Paper prepared for the annual meeting of the Population Association of America, Boston, 1–3 April 2004. http://paa2004.princeton.edu/download.asp?submissionId=41245 [Accessed 19 June 2005].

Zaidi A, Rake K. 2001. *Dynamic microsimulation models: a review and some lessons for SAGE*. SAGE Discussion Paper no. 2. ESRC SAGE Research Group, London School of Economics. http://www.lse.ac.uk/collections/SAGE/pdf/SAGE_DP2.pdf [Accessed 19 June 2005].

Zavodny M. 1999. Determinants of recent immigrants' locational choices. *International Migration Review* **33**: 1014–1030.

Zeng Y, Vaupel JW, Wang Z. 1997. A multi-dimensional model for projecting family households – with an illustrative numerical application. *Mathematical Population Studies* **6**: 187–216.

Zeng Y, Vaupel JW, Zhenglian W. 1999. Household projection using conventional demographic data. In *Frontiers of Population Forecasting*, Lutz W, Vaupel JW, Ahlburg DA (eds). Population Council: New York; 59–87. Supplement to *Population and Development Review* **24**(4).

Zeng Y, Stallard E, Wang Z. 2005. Estimating time-varying age-sex-specific o/e rates of marital status transitions in family household projection or simulation. *Demographic Research* **11**, article 10. http://www.demographic-research.org/ [Accessed 19 June 2005].